Supplementary Information

Supplementary Tables

Supplementary Table 1 – Additional information including nominal capacity (Q_n), preset end of life (EOL), total amount of cells used, and overall average battery life for all datasets. The codename/filename of cells used in train and test set(s) are listed below with their cycle lives.

SNL-NMC

Q_n = 3.0 Ah	2.0-4.2 V	EOL = 80%	Total = 21 cells	Average life = 541 cycles
"SNL_18650_NMC_"		File name – cycle life		
Train set Total: 11 cells Average: 515 cycles		15C_0-100_0.5-1C_a – 164		25C_0-100_0.5-3C_b – 421
		15C_0-100_0.5-2C_a – 156		25C_0-100_0.5-3C_d – 902
		25C_0-100_0.5-0.5C_a – 546		35C_0-100_0.5-1C_b – 523
		25C_0-100_0.5-1C_a – 515		35C_0-100_0.5-1C_d – 653
		25C_0-100_0.5-1C_c – 391		35C_0-100_0.5-2C_b – 765
		25C_0-100_0.5-2C_a – 626		
Test set Total: 10 cells Average: 569 cycles		15C_0-100_0.5-1C_b – 214		25C_0-100_0.5-3C_a – 723
		15C_0-100_0.5-2C_b – 151		25C_0-100_0.5-3C_c – 692
		25C_0-100_0.5-0.5C_b – 569		35C_0-100_0.5-1C_a – 657
		25C_0-100_0.5-1C_b – 779		35C_0-100_0.5-1C_c – 778
		25C_0-100_0.5-1C_d – 467		35C_0-100_0.5-2C_a – 660

SNL-NCA

Q_n = 3.2 Ah	2.5-4.2 V	EOL = 80%	Total = 16 cells	Average life = 435 cycles
"SNL_18650_NCA_"		File name – cycle life		
Train set Total: 9 cells Average: 431 cycles		15C_0-100_0.5-1C_a – 342		25C_0-100_0.5-2C_a – 455
		15C_0-100_0.5-2C_a – 312		35C_0-100_0.5-1C_a – 449
		25C_0-100_0.5-0.5C_a – 383		35C_0-100_0.5-1C_d – 647
		25C_0-100_0.5-1C_a – 493		35C_0-100_0.5-2C_b – 422
		25C_0-100_0.5-1C_c – 378		
Test set Total: 7 cells Average: 439 cycles		25C_0-100_0.5-1C_d – 541		15C_0-100_0.5-2C_b – 329
		25C_0-100_0.5-2C_b – 452		25C_0-100_0.5-0.5C_b – 331
		35C_0-100_0.5-1C_b – 459		25C_0-100_0.5-1C_b – 493
		35C_0-100_0.5-2C_a – 467		

SNL-LFP

Q_n = 1.1 Ah	2.0-3.6 V	EOL = 90%	Total = 19 cells	Average life = 2795 cycles
"SNL_18650_LFP_"		File name – cycle life		
Train set Total: 10 cells Average: 2750 cycles		15C_0-100_0.5-1C_a – 4336		25C_0-100_0.5-3C_a – 2326
		15C_0-100_0.5-2C_a – 3614		25C_0-100_0.5-3C_c – 1581
		25C_0-100_0.5-0.5C_a – 3047		35C_0-100_0.5-1C_a – 1898
		25C_0-100_0.5-1C_b – 3544		35C_0-100_0.5-1C_d – 2326

	25C_0-100_0.5-1C_d – 3026	35C_0-100_0.5-2C_b – 1804
Test set Total: 9 cells Average: 2844 cycles	15C_0-100_0.5-1C_b – 3549 15C_0-100_0.5-2C_b – 3485 25C_0-100_0.5-1C_a – 3515 25C_0-100_0.5-1C_c – 3032 25C_0-100_0.5-2C_a – 3538	25C_0-100_0.5-3C_b – 2147 25C_0-100_0.5-3C_d – 2175 35C_0-100_0.5-1C_b – 1833 35C_0-100_0.5-2C_a – 2326

UL-NCA

Q_n = 3.4 Ah	2.7-4.2 V	EOL = 85%	Total = 21 cells	Average life = 302 cycles
"UL-PUR_18650_NCA_23C_"		File name – cycle life		
Train set Total: 11 cells Average: 300 cycles		N10-EX9_0-100_0.5-0.5C_i – 197 N10-OV8_0-100_0.5-0.5C_h – 168 N15-NA10_0-100_0.5-0.5C_j – 266 N20-EX2_0-100_0.5-0.5C_b – 289 N20-NA6_0-100_0.5-0.5C_f – 199 R10-EX6_2.5-96.5_0.5-0.5C_f – 304		R10-OV5_2.5-96.5_0.5-0.5C_e – 256 R15-NA10_2.5-96.5_0.5-0.5C_j – 610 R20-EX2_2.5-96.5_0.5-0.5C_b – 257 R20-NA8_2.5-96.5_0.5-0.5C_h – 490 R20-OV1_2.5-96.5_0.5-0.5C_a – 268
Test set Total: 10 cells Average: 304 cells		N10-NA7_0-100_0.5-0.5C_g – 174 N15-EX4_0-100_0.5-0.5C_d – 172 N15-OV3_0-100_0.5-0.5C_c – 168 N20-NA5_0-100_0.5-0.5C_e – 162 N20-OV1_0-100_0.5-0.5C_a – 270		R10-NA11_2.5-96.5_0.5-0.5C_k – 449 R15-EX4_2.5-96.5_0.5-0.5C_d – 364 R15-OV3_2.5-96.5_0.5-0.5C_c – 406 R20-NA7_2.5-96.5_0.5-0.5C_g – 363 R20-NA9_2.5-96.5_0.5-0.5C_i – 508

XJTU

Q_n = 2.0 Ah	2.5-4.2 V	EOL = 80%	Total = 23 cells	Average life = 303 cycles
		Code name – cycle life		
Train set Total: 15 cells Average: 249 cycles		3C_battery-1 – 299 3C_battery-3 – 286 3C_battery-5 – 297 3C_battery-7 – 319 3C_battery-9 – 287 3C_battery-11 – 131 3C_battery-13 – 226 3C_battery-15 – 168		3C_battery-2 – 292 3C_battery-4 – 322 3C_battery-6 – 322 3C_battery-8 – 270 3C_battery-10 – 164 3C_battery-12 – 212 3C_battery-14 – 147
Secondary test set Total: 8 cells Average: 402 cells		2C_battery-1 – 390 2C_battery-3 – 393 2C_battery-5 – 403 2C_battery-7 – 402		2C_battery-2 – 407 2C_battery-4 – 396 2C_battery-6 – 408 2C_battery-8 – 420

TRI

Q_n = 1.1 Ah	2.0-3.6 V	EOL = 80%	Total = 123 cells	Average life = 804 cycles
		Code name – cycle life		

Train set Total: 41 cells Average: 674 cycles	b1c1 – 2160	b1c28 – 860	b2c2 – 438	b2c27 – 468
	b1c3 – 1434	b1c30 – 709	b2c4 – 444	b2c29 – 498
	b1c5 – 1074	b1c32 – 731	b2c6 – 511	b2c31 – 492
	b1c7 – 870	b1c34 – 742	b2c11 – 477	b2c33 – 520
	b1c11 – 788	b1c36 – 704	b2c13 – 483	b2c35 – 463
	b1c15 – 719	b1c38 – 617	b2c17 – 494	b2c37 – 478
	b1c17 – 857	b1c40 – 966	b2c19 – 461	b2c39 – 459
	b1c19 – 788	b1c42 – 702	b2c21 – 489	b2c41 – 429
	b1c21 – 559	b1c44 – 616	b2c23 – 527	b2c43 – 462
	b1c24 – 1017	b2c0 – 300	b2c25 – 461	b2c45 – 487
	b1c26 – 870			
Primary test set Total: 42 cells Average: 723 cycles	b1c0 – 1852	b1c27 – 842	b2c5 – 480	b2c30 – 481
	b1c2 – 2237	b1c29 – 917	b2c10 – 561	b2c32 – 519
	b1c4 – 1709	b1c31 – 876	b2c12 – 458	b2c34 – 499
	b1c6 – 636	b1c33 – 757	b2c14 – 485	b2c36 – 535
	b1c9 – 1054	b1c35 – 703	b2c18 – 487	b2c38 – 465
	b1c14 – 880	b1c37 – 648	b2c20 – 502	b2c40 – 499
	b1c16 – 862	b1c39 – 625	b2c22 – 513	b2c42 – 466
	b1c18 – 691	b1c41 – 1051	b2c24 – 495	b2c44 – 457
	b1c20 – 534	b1c43 – 651	b2c26 – 471	b2c46 – 429
	b1c23 – 1014	b1c45 – 599	b2c28 – 509	b2c47 – 713
	b1c25 – 854	b2c3 – 335		
Secondary test set Total: 40 cells Average: 1022 cycles	b3c0 – 1009	b3c11 – 817	b3c21 – 772	b3c33 – 1284
	b3c1 – 1063	b3c12 – 932	b3c22 – 1002	b3c34 – 1158
	b3c3 – 1115	b3c13 – 816	b3c24 – 825	b3c35 – 1093
	b3c4 – 1048	b3c14 – 858	b3c25 – 989	b3c36 – 923
	b3c5 – 828	b3c15 – 876	b3c26 – 1028	b3c40 – 796
	b3c6 – 667	b3c16 – 1638	b3c27 – 850	b3c41 – 786
	b3c7 – 1836	b3c17 – 1315	b3c28 – 541	b3c42 – 1642
	b3c8 – 828	b3c18 – 1146	b3c29 – 858	b3c43 – 1046
	b3c9 – 1039	b3c19 – 1155	b3c30 – 935	b3c44 – 940
	b3c10 – 1078	b3c20 – 813	b3c31 – 731	b3c45 – 1801

Supplementary Table 2 – List of HIs used as the input for Temperature models presented in **Fig. 4** for SNL-LFP, SNL-NMC, SNL-NCA, UL-NCA and XJTU datasets.

	HI	SNL-NMC (Charge)	SNL-NCA (Charge)	SNL-LFP (Discharge)	UL-NCA (Charge)	XJTU (Charge)
T(V)	Kurtosis		✓	✓		
	Max-Min		✓			
	Maximum			✓	✓	
	Mean	✓		✓		✓
	Minimum	✓	✓			
	Skewness			✓	✓	✓
	Variance		✓		✓	✓
dT/dV	Kurtosis				✓	✓
	Max-Min	✓			✓	
	Maximum	✓			✓	✓
	Mean		✓			✓
	Minimum			✓	✓	
	Skewness	✓		✓	✓	
	Variance					

Supplementary Table 3 – List of HIs used as the input for Temperature and Hybrid models presented in **Fig. 4** for TRI dataset.

<table>
<tr><th rowspan="2"></th><th rowspan="2">HI</th><th colspan="2">TRI-Temperature</th><th colspan="2">TRI-Hybrid</th></tr>
<tr><th>Charge (Q)</th><th>Discharge (V)</th><th>Charge (Q)</th><th>Discharge (V)</th></tr>
<tr><td rowspan="7">T(Q) or T(V)</td><td>Kurtosis</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Max-Min</td><td>✓</td><td></td><td>✓</td><td></td></tr>
<tr><td>Maximum</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Mean</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Minimum</td><td>✓</td><td></td><td></td><td></td></tr>
<tr><td>Skewness</td><td></td><td></td><td>✓</td><td></td></tr>
<tr><td>Variance</td><td></td><td>✓</td><td>✓</td><td></td></tr>
<tr><td rowspan="7">dT/dQ or dT/dV</td><td>Kurtosis</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Max-Min</td><td></td><td></td><td></td><td>✓</td></tr>
<tr><td>Maximum</td><td>✓</td><td></td><td>✓</td><td></td></tr>
<tr><td>Mean</td><td>✓</td><td></td><td></td><td></td></tr>
<tr><td>Minimum</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Skewness</td><td>✓</td><td></td><td></td><td></td></tr>
<tr><td>Variance</td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="4">Severson et al.[1]</td><td>Average charge time $_{\text{first 5 cycles}}$</td><td colspan="2" rowspan="4"></td><td colspan="2">✓</td></tr>
<tr><td>Internal resistance $_{\text{cycle 2}}$</td><td colspan="2">✓</td></tr>
<tr><td>Minimum temperature $_{\text{cycle 2-10}}$</td><td colspan="2">✓</td></tr>
<tr><td>Temperature integral $_{\text{cycle 2-10}}$</td><td colspan="2">✓</td></tr>
</table>

Supplementary Table 4 – Benchmark models: Discharge capacity (Q_d) at cycle 2 or 5, and average cycle life of the training set (Train set mean) are used as the Naïve univariate models. Features used in "*Variance*", "*Discharge*", and "*Full*" models from Severson et al.[1] are adapted to use data from the first 10 cycles, instead of the original 100 cycles.

SNL-NMC

	MAE (MAPE%)		RMSE (RMSPE%)	
	Train	Test	Train	Test
$Q_{d\ cycle\ 5}$	122 (24.8)	128 (28.6)	146 (29.0)	147 (40.6)
Train set mean (515)	169 (56.3)	197 (56.7)	218 (97.2)	220 (90.9)
Variance	167 (56.7)	194 (56.8)	214 (98.4)	220 (92.9)
Discharge	73 (16.5)	186 (49.1)	86 (20.4)	257 (87.1)

SNL-NCA

	MAE (MAPE%)		RMSE (RMSPE%)	
	Train	Test	Train	Test
$Q_{d\ cycle\ 5}$	72 (17.2)	69 (16.7)	92 (21.9)	79 (20.1)
Train set mean (431)	71 (16.5)	66 (16.1)	94 (20.6)	75 (19.2)
Variance	68 (15.9)	67 (16.2)	89 (20.5)	76 (19.0)
Discharge	59 (14.1)	71 (16.9)	72 (17.5)	78 (18.6)

SNL-LFP

	MAE (MAPE%)		RMSE (RMSPE%)	
	Train	Test	Train	Test
$Q_{d\ cycle\ 5}$	755 (31.0)	585 (22.3)	842 (37.8)	620 (25.5)
Train set mean (2750)	763 (31.0)	654 (24.4)	860 (36.7)	681 (26.5)
Variance	695 (28.5)	954 (35.3)	798 (34.2)	1021 (37.7)
Discharge	305 (10.4)	814 (31.9)	392 (12.5)	948 (38.9)

UL-NCA

	MAE (MAPE%)		RMSE (RMSPE%)	
	Train	Test	Train	Test
$Q_{d\ cycle\ 5}$	84 (27.4)	47 (17.9)	102 (31.4)	56 (22.9)
Train set mean (300)	91 (30.5)	114 (45.7)	127 (38.8)	124 (53.4)
Variance	91 (30.5)	114 (45.7)	127 (38.8)	124 (53.4)
Discharge	81 (26.2)	65 (26.3)	101 (28.6)	74 (31.7)

XJTU

	MAE (MAPE%)		RMSE (RMSPE%)	
	Train	Test	Train	Test
$Q_{d\ cycle\ 2}$	30 (14.5)	84 (21.0)	38 (19.8)	92 (22.9)
Train set mean (249)	60 (29.1)	153 (38.0)	66 (37.5)	153 (38.0)
Variance	58 (28.2)	105 (26.2)	64 (35.7)	118 (29.7)
Discharge	25 (12.0)	90 (22.2)	33 (16.3)	103 (25.2)

TRI

	MAE (MAPE%)			RMSE (RMSPE%)		
	Train	Primary Test	Secondary Test	Train	Primary Test	Secondary Test
$Q_{d\ cycle\ 2}$	192 (26.4)	229 (26.9)	482 (44.1)	315 (32.0)	388 (31.9)	557 (45.8)
Train set mean (674)	221 (33.5)	248 (31.9)	355 (30.9)	323 (40.0)	389 (37.6)	450 (34.2)
Variance	193 (25.9)	240 (29.4)	410 (36.5)	317 (31.1)	397 (36.8)	496 (39.2)
Discharge	182 (24.4)	239 (29.0)	466 (42.3)	306 (29.3)	401 (35.8)	543 (44.2)
Full	80 (10.5)	122 (14.5)	304 (26.8)	143 (15.3)	202 (18.1)	380 (29.7)

Supplementary Table 5 – Exploring the optimal set of HIs for Temperature model on the TRI dataset using the average HI of the first 3 to 10 cycles, evaluated on the test sets where MAE and RMSE of the primary and secondary test sets are provided below. Note that the first initialization cycle was excluded for all average ranges.

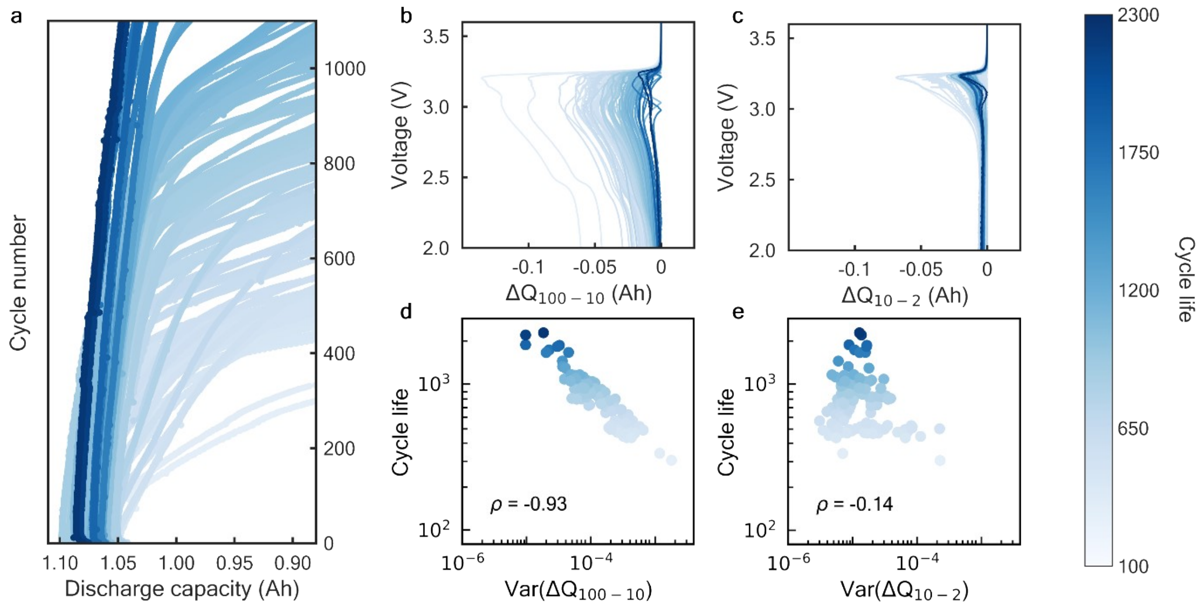
	HI	Avg_3		Avg_4		Avg_5		Avg_6		Avg_7		Avg_8		Avg_9		Avg_{10}	
		Ch (Q)	Dis (V)	Ch (Q)	Ch (Q)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)
T(Q) or T(V)	Kurtosis																
	Max-Min	✓		✓	✓	✓		✓		✓	✓	✓	✓	✓		✓	
	Maximum																
	Mean																
	Minimum															✓	
	Skewness	✓				✓											
	Variance		✓		✓		✓		✓		✓		✓		✓		✓
dT/dQ or dT/dV	Kurtosis																
	Max-Min										✓		✓				
	Maximum	✓		✓		✓		✓		✓		✓		✓		✓	
	Mean					✓								✓		✓	
	Minimum																
	Skewness			✓		✓		✓		✓		✓		✓		✓	
	Variance					✓		✓		✓	✓	✓	✓	✓			
Primary test	MAE	155		155		139		130		120		126		128		129	
	RMSE	238		257		221		205		202		222		221		216	
Secondary test	MAE	198		205		197		182		180		183		176		174	
	RMSE	264		276		269		259		256		266		264		257	

Supplementary Table 6 – Exploring the optimal set of HIs for Hybrid model on the TRI dataset using the average HI of the first 3 to 10 cycles, evaluated on the test sets where MAE and RMSE of the primary and secondary test sets are provided below. Note that the first initialization cycle was excluded for all average ranges.

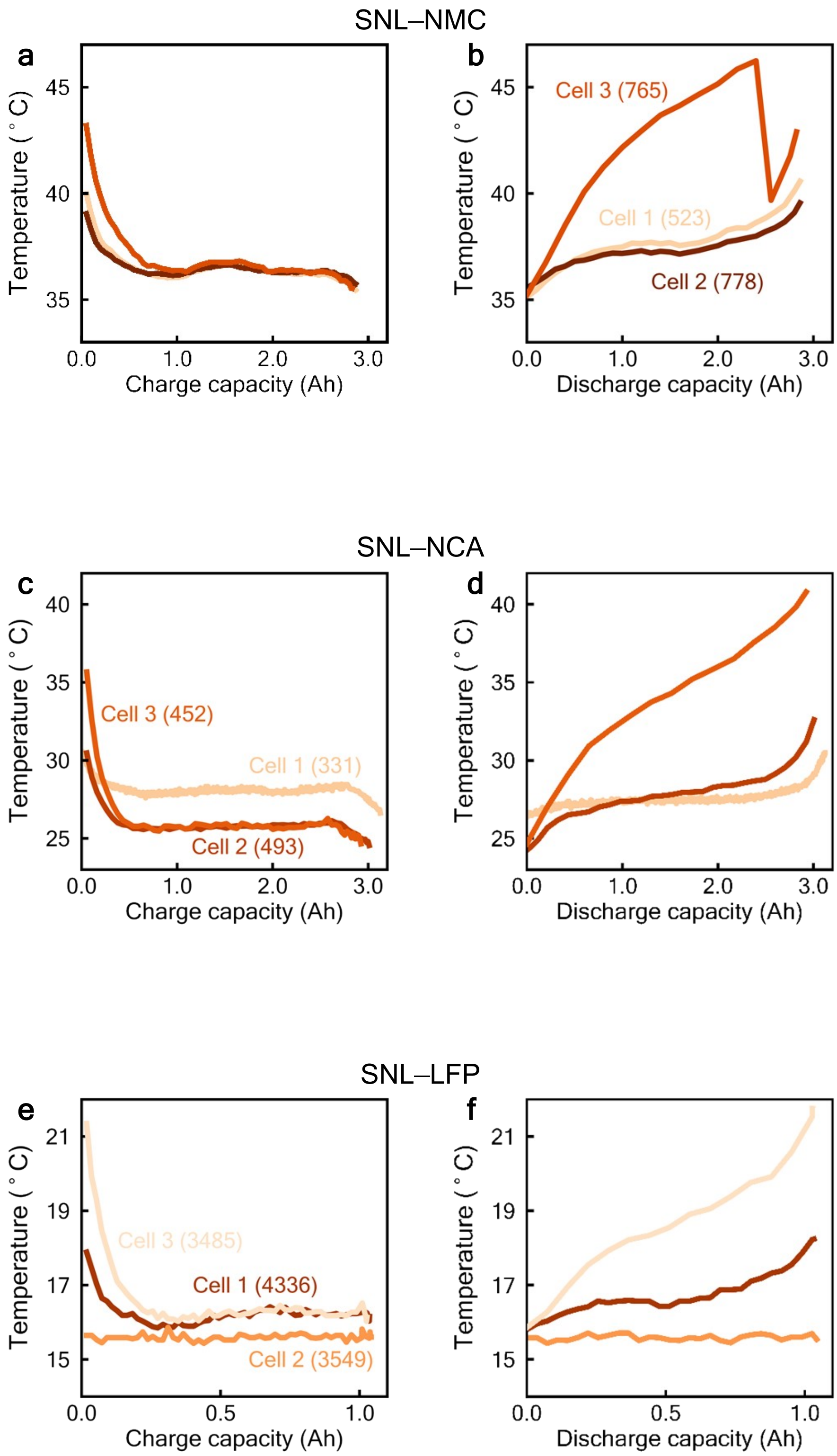
	HI	Avg_3		Avg_4		Avg_5		Avg_6		Avg_7		Avg_8		Avg_9		Avg_{10}	
		Ch (Q)	Dis (V)	Ch (Q)	Ch (Q)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)	Ch (Q)	Dis (V)
T(Q) or T(V)	Kurtosis																
	Max-Min	✓			✓		✓	✓		✓		✓		✓		✓	
	Maximum					✓											
	Mean																
	Minimum																
	Skewness	✓		✓		✓		✓		✓		✓		✓		✓	
	Variance	✓	✓		✓	✓	✓	✓		✓	✓	✓		✓		✓	
dT/dQ or dT/dV	Kurtosis																
	Max-Min								✓				✓		✓		✓
	Maximum	✓		✓		✓		✓		✓		✓		✓		✓	
	Mean																
	Minimum																
	Skewness																
	Variance																
Severson et al.[1]	Average charge time first 5 or x cycles	✓		✓		✓		✓		✓		✓		✓		✓	
	Internal resistance cycle 2							✓				✓		✓		✓	
	Minimum temperature cycle 2-x	✓		✓		✓		✓		✓		✓		✓		✓	
	Temperature integral cycle 2-x	✓		✓		✓		✓		✓		✓		✓		✓	
Primary test	MAE	85		96		88		88		89		94		95		94	
	RMSE	135		141		138		135		138		145		147		146	
Secondary test	MAE	153		157		150		160		151		148		148		148	
	RMSE	211		202		204		204		204		204		202		203	

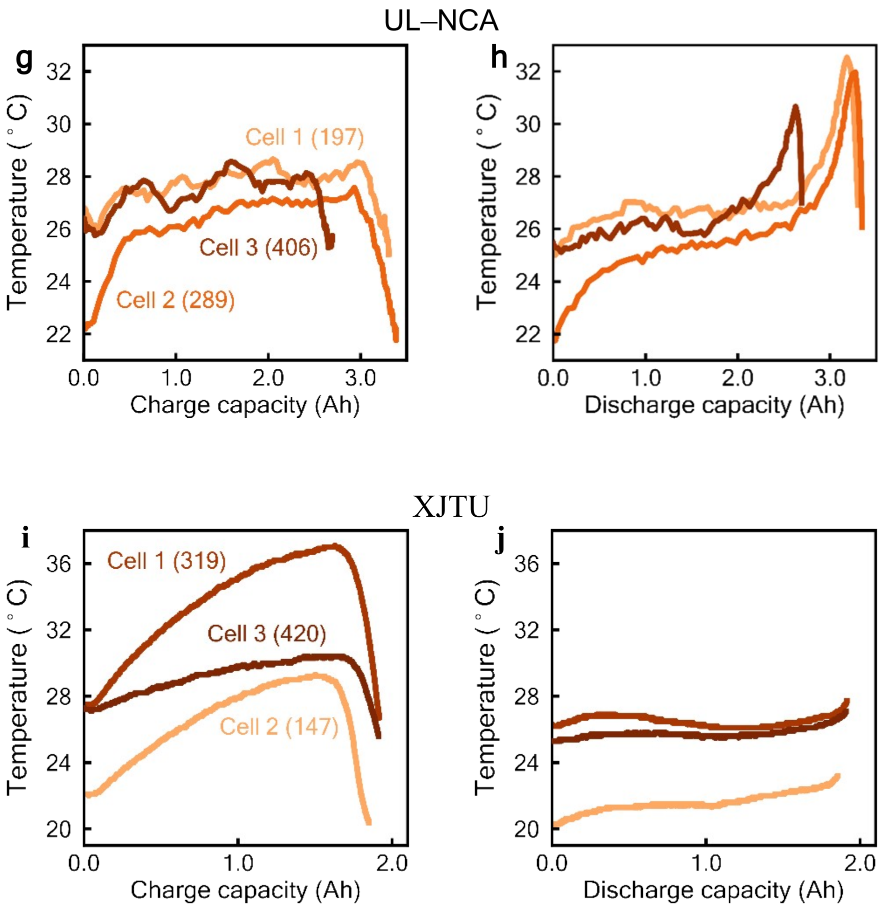
Supplementary Figures

Supplementary Figure 1 – Visualization of TRI dataset distribution and the highlighted feature for cycles difference of 100-10 and 10-2. **a,** Cycle life as a function of discharge capacity until EOL, 0.88 Ah. **b-c,** Plot of discharge capacity curve difference for **b,** cycle 100 – 10 ($\Delta Q_{100\text{-}10}(V)$), and **c,** cycle 10-2 ($\Delta Q_{10\text{-}2}(V)$), as a function of voltage. **d-e,** The feature variance of $\Delta Q_{y\text{-}x}(V)$ is plotted against cycle life for **d,** $\Delta Q_{100\text{-}10}(V)$, and **e,** $\Delta Q_{10\text{-}2}(V)$ with a Pearson correlation coefficient of -0.93 and -0.14, respectively, under a logarithmic-scale on both axes. In all figures, color opacity indicates the EOL of cells.

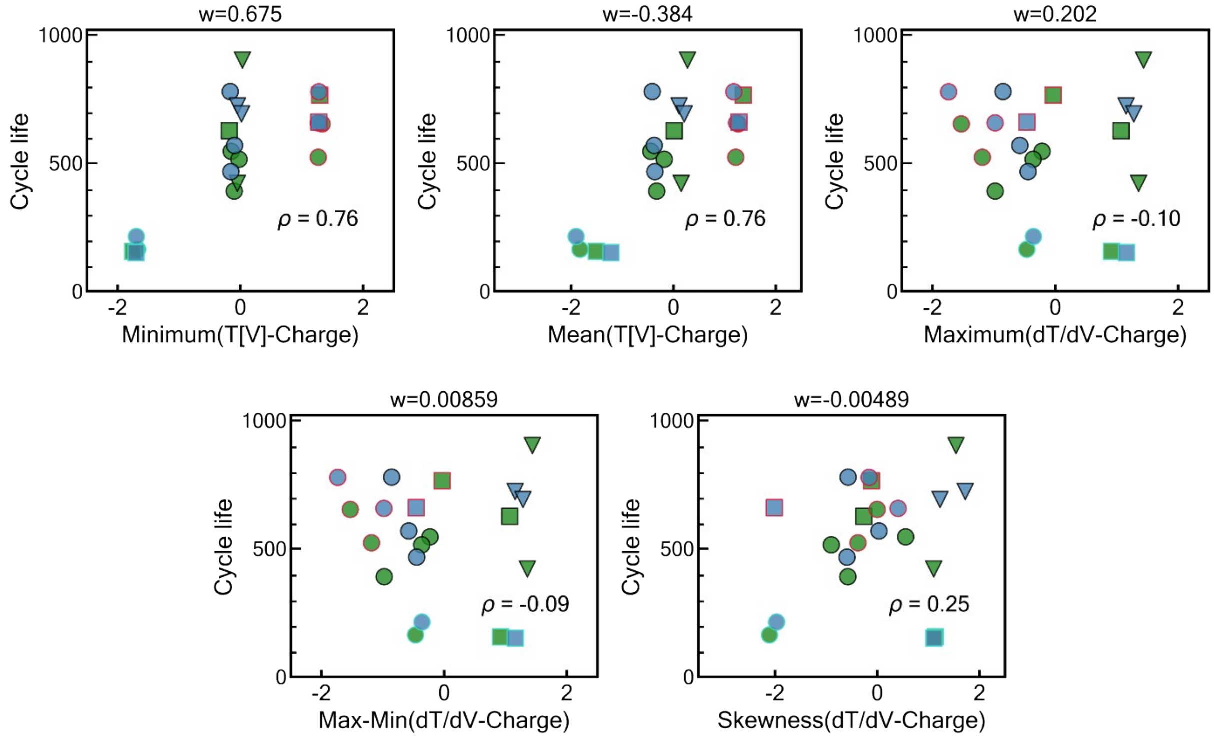


Supplementary Figure 2 – Cycling temperature profiles of three sample cells at cycle 10 with cycle life annotated in the bracket, and (dis)charge capacity as the x-axis for SNL, UL and XJTU datasets.

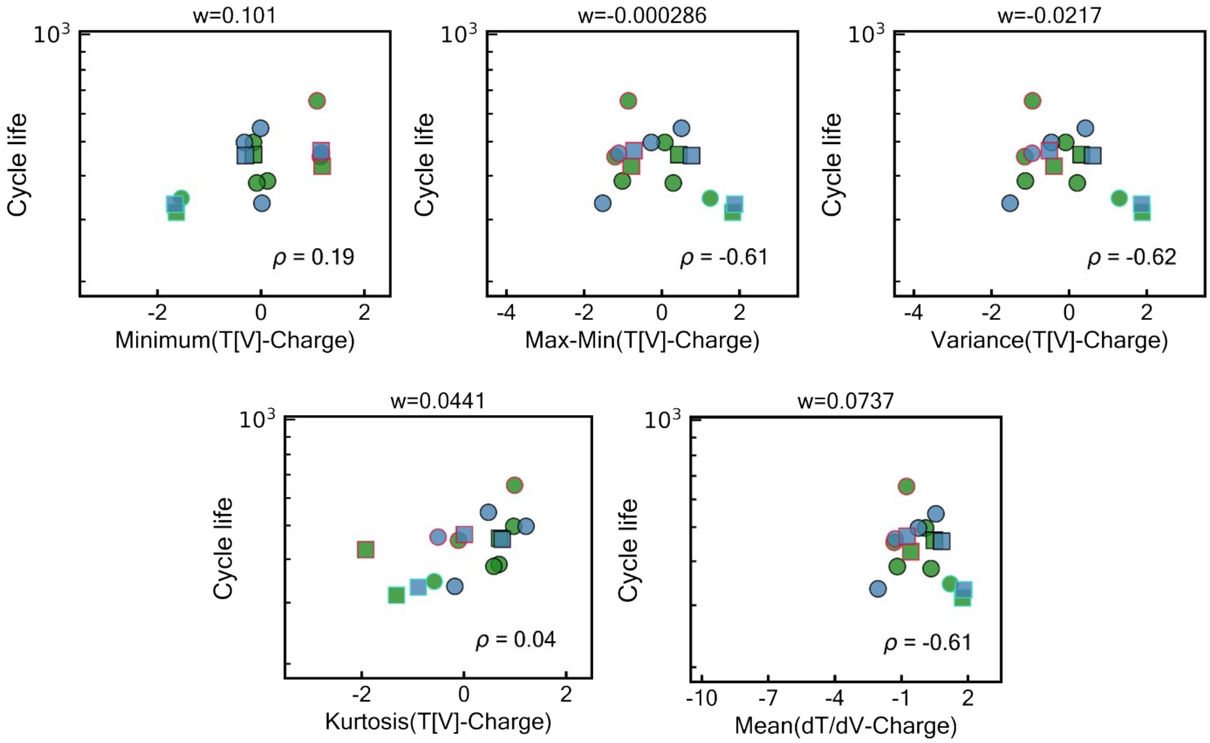



UL–NCA
g
h
Cell 1 (197)
Cell 3 (406)
Cell 2 (289)
Temperature (° C)
Charge capacity (Ah)
Discharge capacity (Ah)
XJTU
i
j
Cell 1 (319)
Cell 3 (420)
Cell 2 (147)
Temperature (° C)
Charge capacity (Ah)
Discharge capacity (Ah)

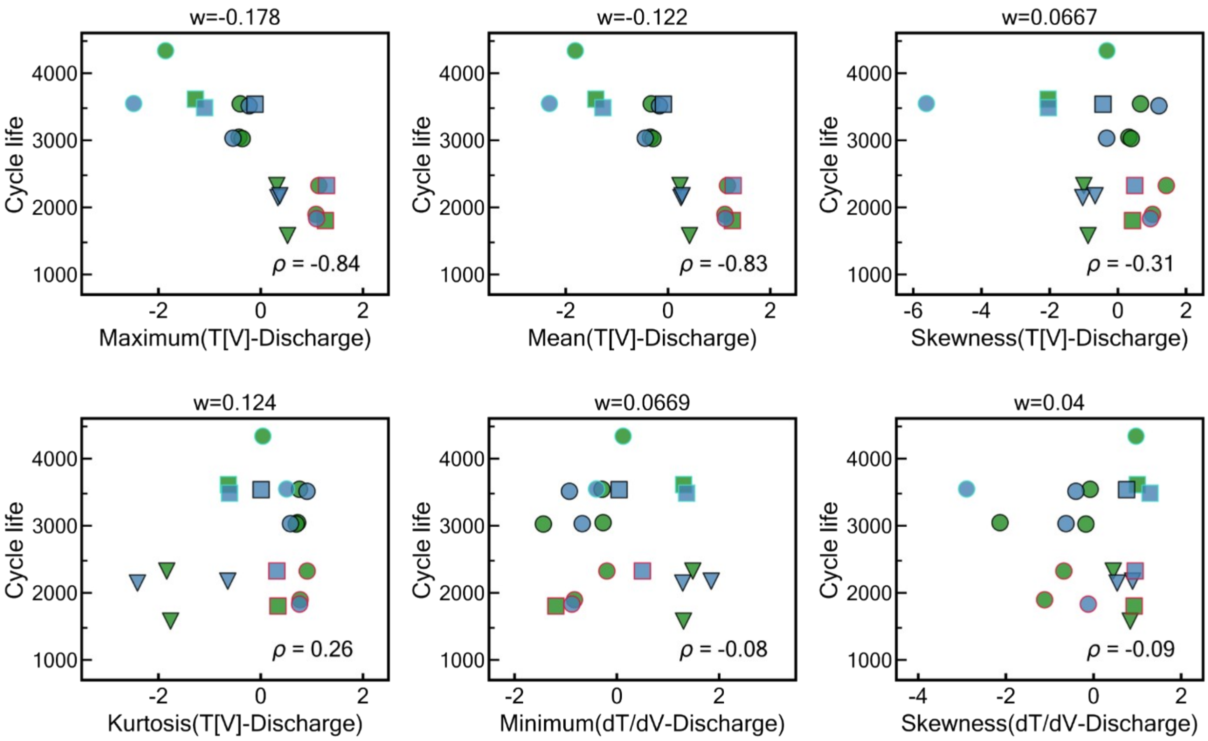
Supplementary Figure 3 – Five HIs used in the Temperature model for SNL-NMC were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under a normal scale on the y-axis. The marker colors represent the train (green) and test (blue) data, the marker shape represents the discharge C-rates (circle: 1 C, square: 2C, triangle: 3C), and the marker outline color represents the environment temperature (light blue: 15ºC, black: 25ºC, red: 35ºC). The coefficient (w) is notated at the top of each HI plot.



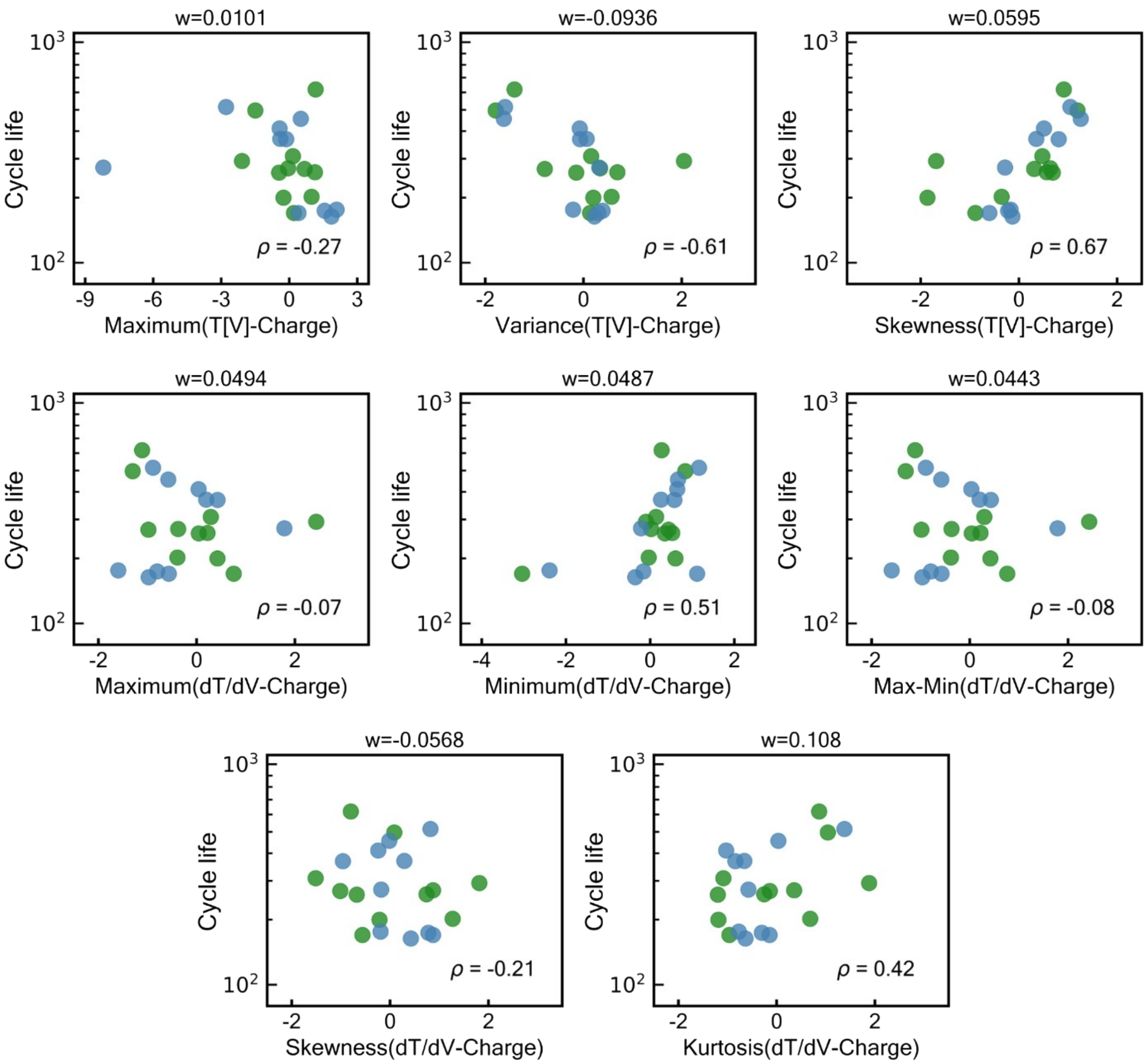
Supplementary Figure 4 – Five HIs used in the Temperature model for SNL-NCA were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under a logarithmic scale on the y-axis. The marker colors represent the train (green) and test (blue) data, the marker shape represents the discharge C-rates (circle: 1 C, square: 2C), and the marker outline color represents the environment temperature (light blue: 15ºC, black: 25ºC, red: 35ºC). The coefficient (w) is notated at the top of each HI plot.



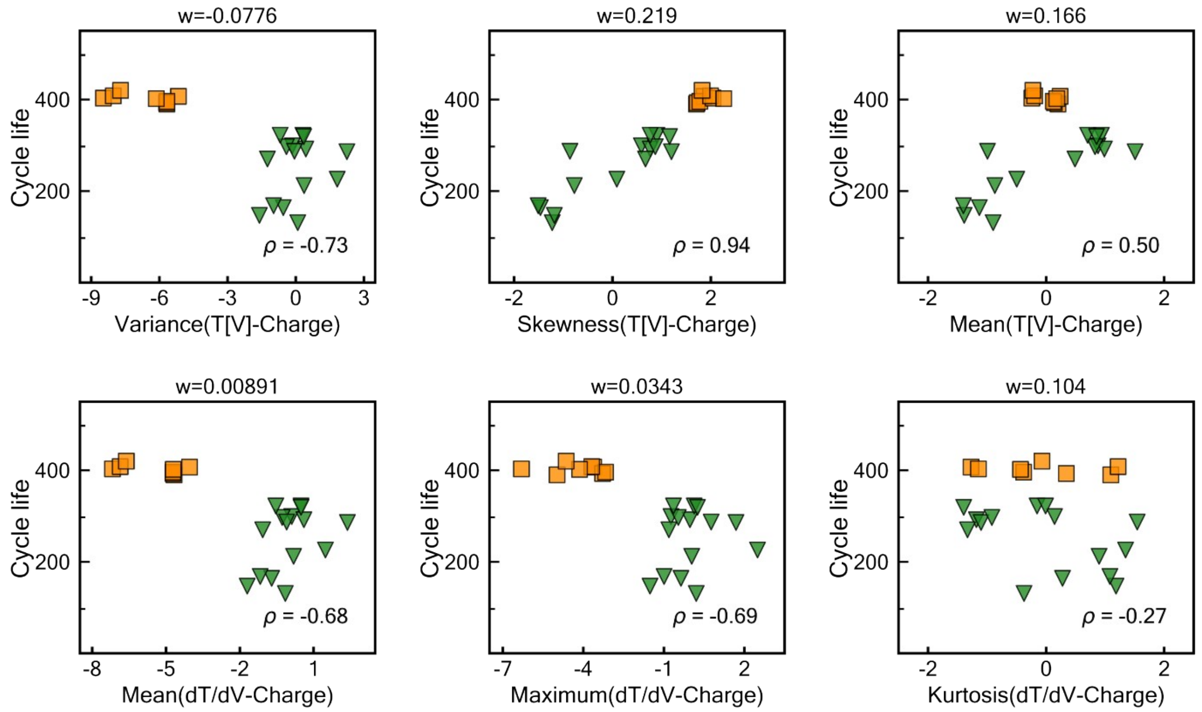
Supplementary Figure 5 – Six HIs used in the Temperature model for SNL-LFP were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under normal scale on the y-axis. The marker colors represent the train (green) and test (blue) data, the marker shape represents the discharge C-rates (circle: 1 C, square: 2C, triangle: 3C), and the marker outline color represents the environment temperature (light blue: 15ºC, black: 25ºC, red: 35ºC). The coefficient (w) is notated at the top of each HI plot.



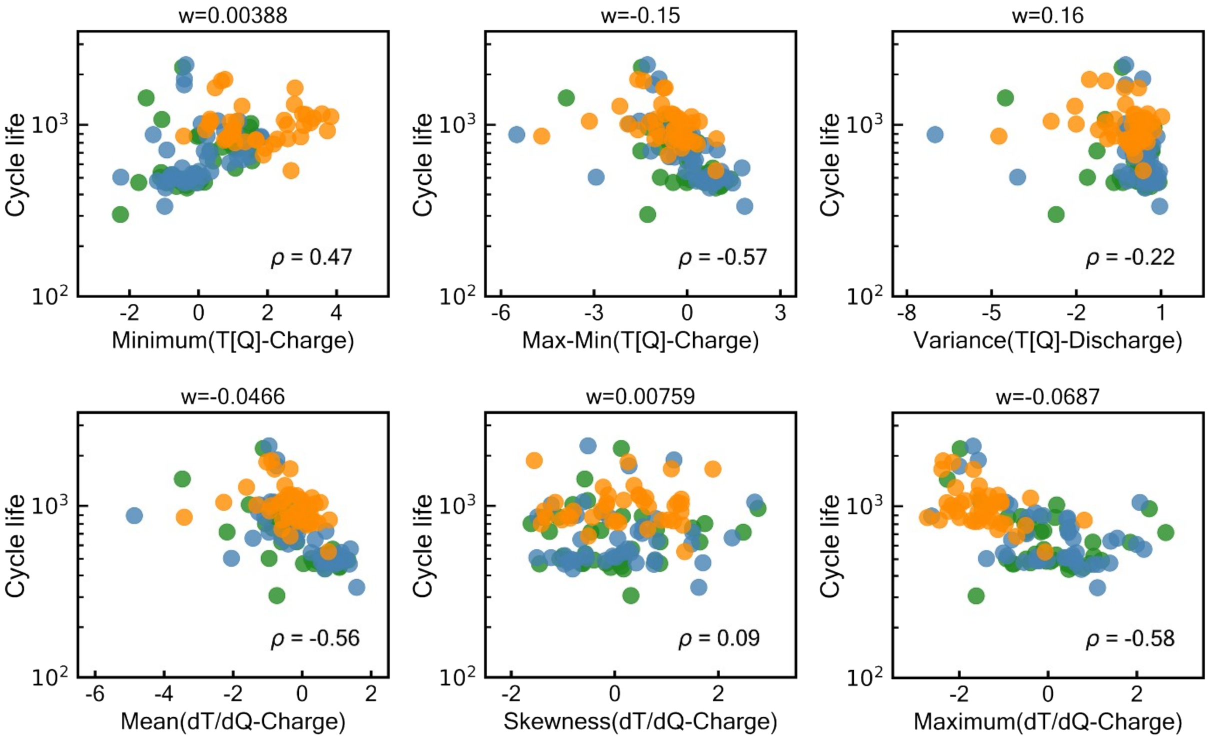
Supplementary Figure 6 – Eight HIs used in the Temperature model for UL-NCA were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under a logarithmic scale on the y-axis. The marker colors represent the train (green) and test (blue) data. The coefficient (w) is notated at the top of each HI plot.



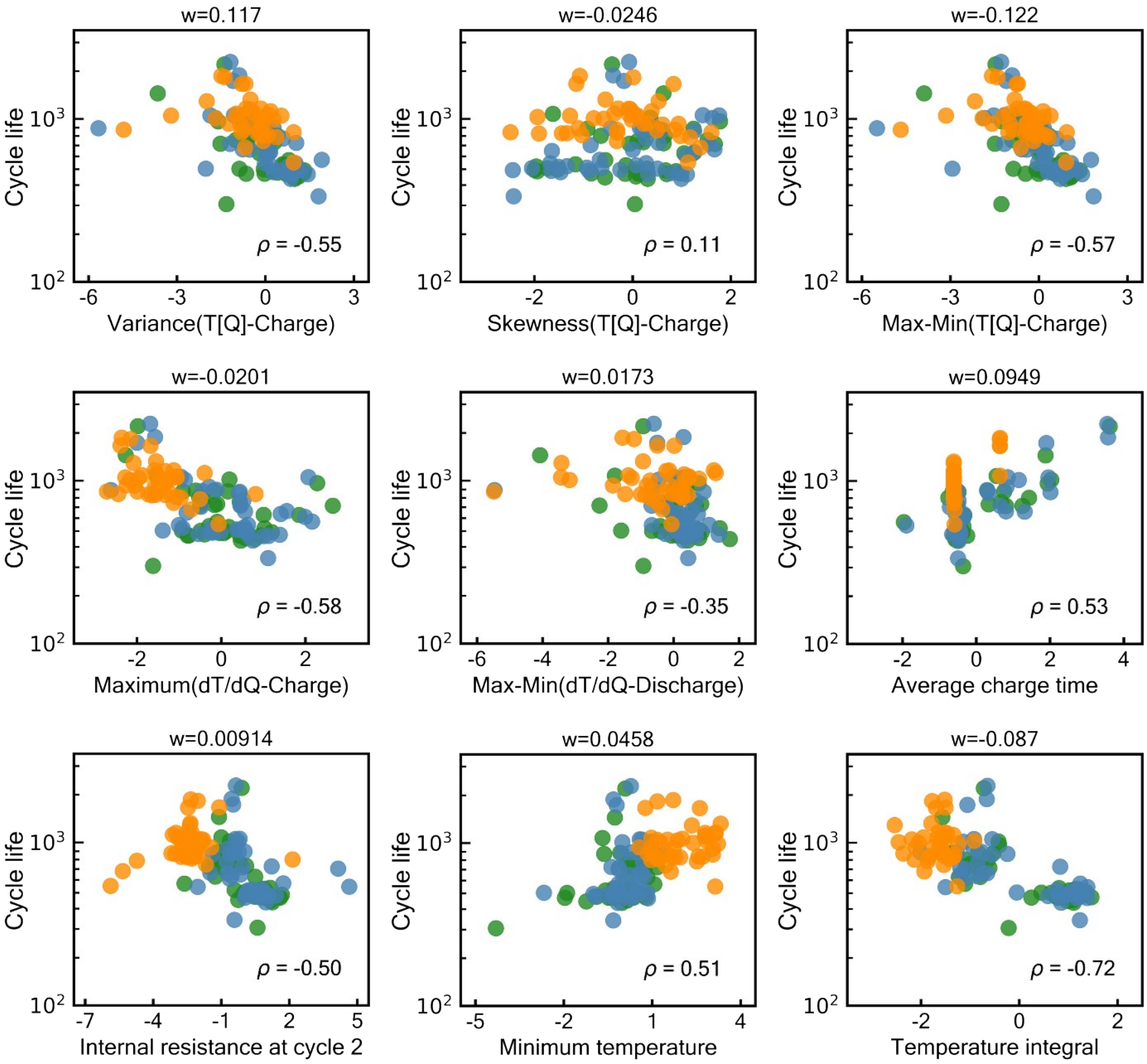
Supplementary Figure 7 – Six HIs used in the Temperature model for XJTU were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under normal scale on the y-axis. The marker colors represent the train (green) and test (blue) data, and the marker shape represents the charge C-rates (square: 2C, triangle: 3C). The coefficient (w) is notated at the top of each HI plot.



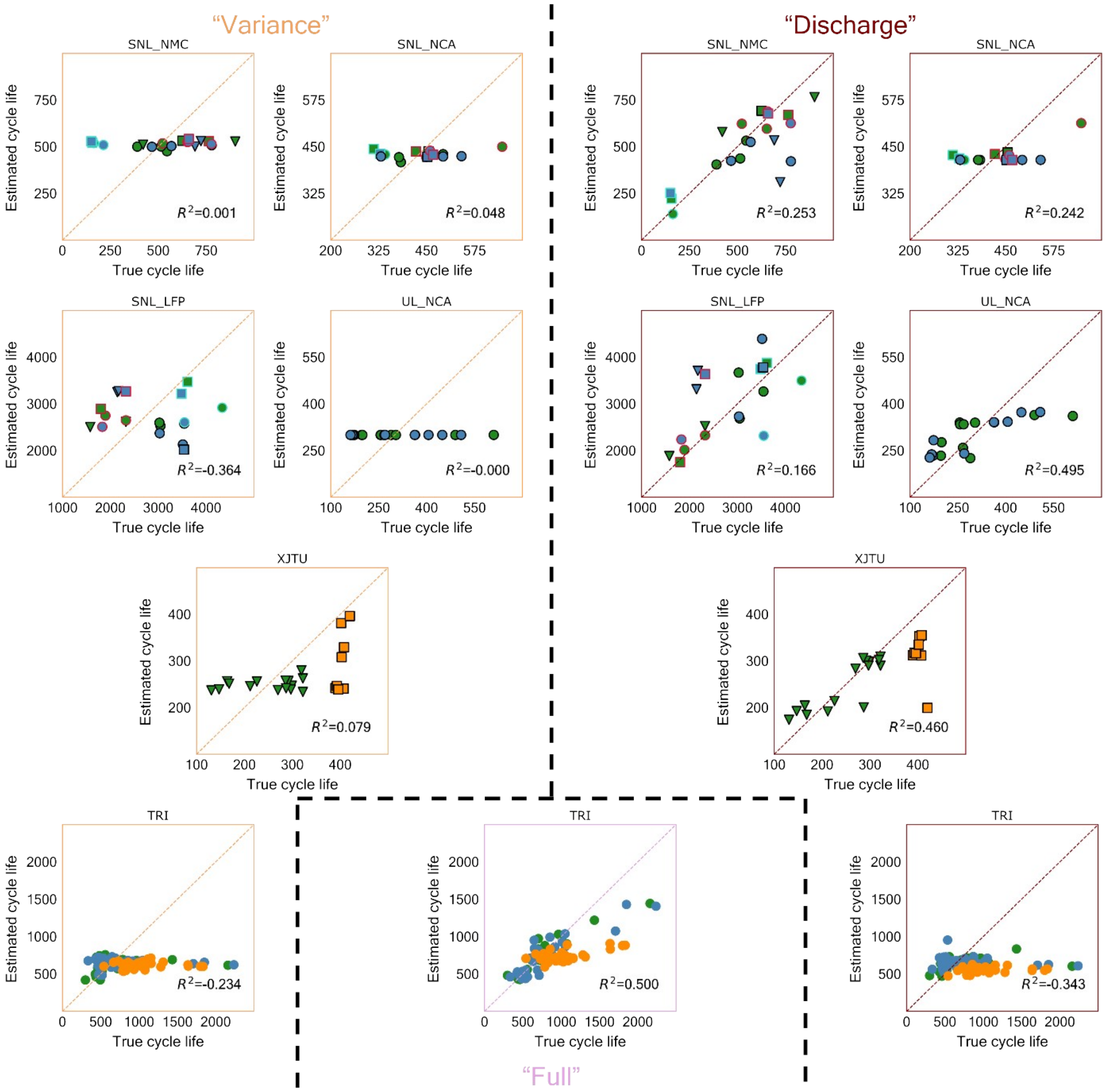
Supplementary Figure 8 – Six HIs used in the Temperature model for TRI were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under a logarithmic scale on the y-axis. The marker colors represent the train (green), primary test (blue), and secondary test (orange) data. The coefficient (w) is notated at the top of each HI plot.



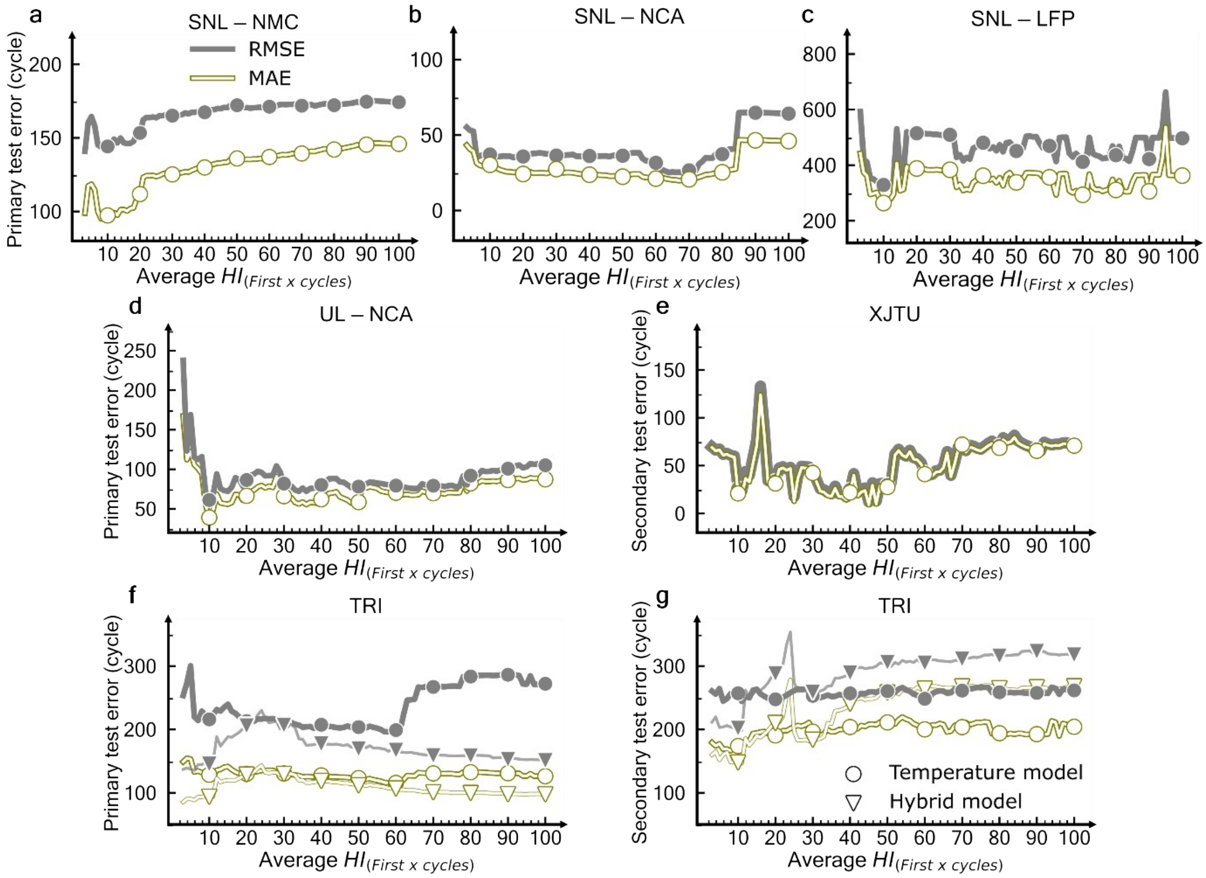
Supplementary Figure 9 – Nine HIs used in the Hybrid model for TRI were preprocessed to their standardized values (i.e. z-score) as shown on the x-axis based on their training set, plotted against cycle life under logarithmic scale on the y-axis. The marker colors represent the train (green), primary test (blue), and secondary test (orange) data. The coefficient (w) is notated at the top of each HI plot.



Supplementary Figure 10 – "*Variance*", "*Discharge*" and "*Full*" models, recreated using the first 10 cycles on each dataset, were used as benchmarks. The models are indicated by the plot line colors orange, red, and violet respectively, and the related dataset name is notated at the top of each plot. Only the TRI dataset uses the "*Full*" model as internal resistance information was unavailable on other datasets.



Supplementary Figure 11 – Performance of Temperature and Hybrid models trained using the average HI across different cycle ranges from 3 to 100 cycles. MAE (solid gray line) and RMSE (hollow olive line) of the trained Average HI_x cycles models are evaluated against the test sets in each dataset, on the primary test set for **a,** SNL-NMC, **b,** SNL-NCA, **c,** SNL-LFP, **d,** UL-NCA, **f,** TRI; and secondary test set for **e,** XJTU, **g,** TRI. The Temperature model is indicated by circle markers, whereas the Hybrid model is represented using triangle markers. Note that the first initialization cycle was excluded for all average ranges.



Supplementary Notes

Supplementary Note 1 – Online dataset preparation and feature engineering

In this study, we used four online datasets provided by Toyota Research Institute (TRI)[1], Sandia National Laboratories (SNL)[2], Underwriters Laboratories Inc. and Purdue University (UL-PUR)[3], and Xi'an Jiaotong University (XJTU)[4] as shown in **Section 2.2** and **Supplementary Table 1**. The raw data were initially prepared in MATLAB as struct format, then converted into Python dictionaries for preprocessing (e.g. filling in missing data), standardizing the formatting of array vectors, and extracting key features, or Health Indicators (HIs). Following this, the data were transformed and used for model training, during which the HI values were selected and rescaled. To ensure consistent data processing, we linearly interpolated the raw temperature data, focusing specifically on the constant-current (CC) region. This interpolation produced temperature vectors with 100 evenly spaced values as a function of voltage (or capacity), ensuring uniform vector formatting. Since each dataset used different battery chemistries, the cutoff voltages (or capacities) for interpolation were as follows:

- SNL-NMC: Charge 3.15 – 4.195 V; Discharge 4.195 – 2.005 V
- SNL-NCA: Charge 3.5 – 4.195 V; Discharge 4.195 – 2.505 V
- SNL-LFP: Charge 2.995 – 3.595 V; Discharge 3.595 – 2.005 V
- UL-NCA: Charge 2.95 – 4.195 V; Discharge 4.195 – 2.705 V
- TRI: Charge 0 – 0.88 Ah; Discharge 3.6 – 2.04 V
- XJTU: Charge 3.6 – 4.195 V; Discharge 4.15 – 2.5 V

We applied small adjustments to the interpolation cutoff voltages to accommodate overpotential at the start of charging, and to capture the final points before the constant-voltage (CV) phase during discharge. Although we tested including the entire CC-CV region for the TRI dataset, this approach significantly increased the Pearson correlation coefficient for some statistical HIs in both the training and primary test sets. However, to avoid overfitting the TRI training set and sacrificing accuracy on the secondary test set, we ultimately decided not to use the full CC-CV region. The secondary test set for TRI is critical, as it evaluates the generalizability of our proposed temperature-based model.

For the TRI dataset, we used capacity instead of voltage for the charging region interpolation. This decision was made because the charging protocols often involved multi-step current settings at specific states of charge (SOC), leading to sharp voltage fluctuations when the current was switched. These fluctuations introduced noise into the temperature-voltage profile, so using capacity instead of voltage provided a cleaner data set for interpolation. For the other datasets, voltage was used for interpolation to evaluate the variation in ohmic overpotential at the beginning of charge or discharge between battery cells. Additionally, the upper cutoff capacity value for segregating the CC stage varied slightly between cells of the same cycle number, making voltage a more convenient variable for these datasets.

After interpolation, we extracted statistical HIs from the temperature vectors, with their formula expressed in **Supplementary Note 2**. For model prediction, we transformed the target variable (i.e., cycle life) to its $\log_{10}$ value. Severson et al[1] reported a better relationship between their features and $\log_{10}$ cycle life, as it linearizes the observed exponential trend when using the normal scale for cycle life. We applied this $\log_{10}$ transformation to SNL-NCA and UL-NCA, but not SNL-LFP, SNL-NMC and XJTU. In the case of LFP dataset, which has a very high cycle life range (i.e. 1500-4500), the $\log_{10}$ transformation altered the relationship between the HIs and cycle life in a way that was less effective than using the normal scale. Similarly, for SNL-NMC and XJTU, we found that using the normal cycle life scale yielded slightly better or similar results.

Finally, we used the extracted HIs to train our models via ElasticNet regularization, combined with an exhaustive search for feature selection. The Pearson correlation coefficient was employed to rank and prioritize the selection of HIs during the optimization process, reducing training time. We also standardized the features during model training to ensure that all feature weightings were comparable.

Supplementary Note 2 – Formulation of statistical temperature HIs

T[V] or T[Q] is a temperature vector of S = 100 elements with equal intervals of V or Q. The temperature gradient dT/dV or dT/dQ is derived from the gradient of each corresponding temperature vector with length S – 1 (i.e. 99 elements). The statistical HI summary formulas are shown below[4,5] where the vectors T[V], T[Q], dT/dV, and dT/dQ are represented as X interchangeably, $\bar{X}$ is the vector mean, and x_i indicates the element of each vector.

1. Maximum = $\log_{10}(|max(X)|)$
2. Minimum = $\log_{10}(|min(X)|)$
3. Maximum – Minimum = $\log_{10}(|max(X) - min(X)|)$
4. Mean = $\log_{10}\left(\left|\frac{1}{S}\sum_{i=1}^{S} x_i\right|\right)$
5. Variance = $\log_{10}\left(\left|\frac{1}{S}\sum_{i=1}^{S} (x_i - \bar{X})^2\right|\right)$
6. Skewness = $\log_{10}\left(\left|\frac{\frac{1}{S}\sum_{i=1}^{S} (x_i - \bar{X})^3}{\left(\sqrt{\frac{1}{S}\sum_{i=1}^{S} (x_i - \bar{X})^2}\right)^3}\right|\right)$
7. Kurtosis = $\log_{10}\left(\left|\frac{\frac{1}{S}\sum_{i=1}^{S} (x_i - \bar{X})^4}{\left(\frac{1}{S}\sum_{i=1}^{S} (x_i - \bar{X})^2\right)^2}\right|\right)$

Supplementary Note 3 – Limitations of temperature HI efficacy

This study covers various cycling operations, including diverse environmental temperatures and cathode chemistries. However, the current design remains somewhat inflexible and leaves room for improvement, particularly in terms of extracting HIs from partial cycling data (i.e. less than 100% DOD) and selecting an optimal set of universal HIs. For instance, the HIs used in our models were chosen based on the best generalization performance for each individual dataset. As a result, no single set of HIs consistently stands out across all datasets, as their predictive impact varies depending on the specific characteristics of each dataset.

Statistical temperature HIs were computed for each of the first 10 cycles and the average values of each HI were used as the model input (i.e. 1-D vector). We initially explored the performance of the model when using HIs from later cycles, but found that adding more data did not necessarily improve predictive accuracy, as illustrated in **Fig. 5 and Fig. S11**. There are several possible reasons for this result. First, the models were trained on the best set of averaged HIs, but the importance of each HI likely shifts from cycle to cycle. While it would be possible to find the best model for each individual cycle through an exhaustive search, this approach would be time-consuming and was not pursued in this study.

Second, our study does not account for the evolution of the temperature profile between early and later cycles, a factor that was considered by Severson et al.[1] As batteries age, the magnitude of temperature fluctuations tends to increase, leading to larger temperature swings. When extracting HIs from single-cycle data, this consistent shift in temperature magnitude can make differences between cells less apparent. As a result, the Pearson correlation coefficient between the HIs and cycle life remains relatively unchanged across different cycles. Further investigation could involve comparing the distribution of temperature profiles between early and later cycles to gain deeper insights into this phenomenon.

Data quality and quantity play a crucial role in model performance. For example, in the Hybrid model used on the TRI dataset, errors in the secondary test set were higher than those observed in the training and primary test sets. This discrepancy could be indicative of overfitting, which may have originated from either model complexity or insufficient data. In this case, the latter is likely the primary issue, as Severson et al.[1] noted that calendar aging effects persist in their secondary test set. Our model may not fully capture these calendar aging patterns, which could explain the elevated error rates.

Additionally, we aimed to develop accurate early-cycle models by minimizing the training data, which was done via a balanced train-test split ratio. Our Temperature model performed better than expected on the XJTU dataset with a 2:1 split ratio. The train set consists of 3C charging rate protocols, which predicted the test set, containing 2C charging rate (i.e. out-of-protocol predictions), very accurately as shown in **Fig. 4**. However, we could not replicate similar results on the other datasets, UL-NCA and SNL datasets, which are heavily affected by data insufficiency. These datasets are much smaller than TRI, with each model being trained on only around 10 data points for the k-fold splits and a 1:1 split ratio. This limited amount of data combined with more experimental dimensionality including environmental temperature and discharge rate settings may be inadequate for effective feature selection on 14 HIs. While it would be possible to adjust the train-test ratio, a larger dataset would be required for more robust model validation against training data overfitting.

We note that the proposed temperature HIs have been specifically tested on lithium-ion cells, leveraging the heat generated from resistive components and charge transfer processes inherent in intercalation-based systems. While effective for LIBs due to their electrochemical and thermal behaviors, the effectiveness of these HIs may vary for energy storage technologies with different degradation mechanisms, such as conversion-based systems. Further research is needed to validate their universality across various technologies. Additionally, the LIB datasets used in this study feature cylindrical-18650 cells with single-point temperature measurements. While these temperature HIs could potentially be adapted for different cell configurations, incorporating a greater number and strategic placement of temperature sensors may enhance data collection and analysis.

References

[1] K. Severson, P. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. Chen, M. Aykol, P. Herring, D. Fraggedakis and M. Bazant, *Nat. Energy*, 2019, **4**, 383-391.

[2] Y. Preger, H. Barkholtz, A. Fresquez, D. Campbell, B. Juba, J. Romàn-Kustas, S. Ferreira and B. Chalamala, *J. Electrochem. Soc.*, 2020, **167**, 120532.

[3] D. Juarez-Robles, J. Jeevarajan and P. Mukherjee, *J. Electrochem. Soc.*, 2020, **167**, 160510.

[4] F. Wang, Z. Zhai and Z. Zhao, *Nat. Commun.*, 2024, **15**, 4332.

[5] C. Harris, K. Millman and S. van der Walt, *Nature*, 2020, **585**, 357-362.

[6] F. Pedregosa, *J. Mach. Learn. Res.*, 2011, **12**, 2825-2830.